

REMARKS

The application has been amended to include a substitute Sequence Listing in conformity with 37 C.F.R. § § 1.821-1.825. Submitted herewith is the substitute Sequence Listing, a copy of the substitute Sequence Listing in computer readable form, as well as a Statement under 37 C.F.R. § § 1.821(f) and 1.825. The application also has been amended to contain appropriate sequence identifiers in the specification.

It is respectfully submitted that no new matter has been added by the amendment. However, if any question remain after consideration of the instant amendments, the Office is kindly requested to contact applicants' attorney at the address or telephone number given herein.

Respectfully submitted,



Tawni L. Wilhelm
Registration No. 47,456
Attorney for Applicants
TRASKBRITT, PC
P. O. Box 2550
Salt Lake City, Utah 84110-2550
Telephone: (801) 532-1922

Date: May 20, 2002

ACT/TLW/bv

Enclosure:

Sequence Statement with paper copy and CRF disk
Response to Notice to Comply

APPENDIX A

(VERSION WITH MARKINGS TO SHOW CHANGES MADE)

(Serial No. 10/016,516)

[0023] Using the VAST protocol (<http://www.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml>), it is possible to find identical or partly identical structures. VAST uses an algorithm that can predict the presence of identical elements in presubmitted structures with an example template structure. Here, the structure of 1F2X (SEQ ID NO:87), a functional camelid variable domain (VHH) containing an immunoglobular domain, was chosen as a template to identify naturally occurring proteins in all kinds of organisms that share structural elements. Next, the amino acid sequence of some of the retrieved proteins was aligned 1QD0 (SEQ ID NO:8), 8FAB (SEQ ID NO:9), 1VSC (SEQ ID NO:10), 1NS3 (SEQ ID NO:11), 1F97 (SEQ ID NO:12), 1FNF (SEQ ID NO:13), 1CFB (SEQ ID NO:14), 1IAR (SEQ ID NO:15), and 1F4H (SEQ ID NO:16). The location of all 9 beta-elements organized in two beta-sheets is indicated by underlining. Some matched proteins contain all 9 beta-elements while others lack one or more of these. Despite the lack of one or more structural components, such proteins still form a basic common structure. Amino acid sequence comparisons show that there is hardly any conservation although the 3D-structures of these proteins is, at least partially, identical. Structurally identical proteins from all kinds of organisms were retrieved this way, varying from bacteria to flies to human.

[0024] a) Amino acid and corresponding DNA sequences of a primary scaffold in which example CDR regions are inserted. These sets of CDR regions, depicted underlined in the figure, originate from 3 different camelid-derived VHH variable regions, known as 1BZQ (nucleotide sequence SEQ ID NO:19, protein sequence SEQ ID NO:20), 1HCV (nucleotide sequence SEQ ID NO:21, protein sequence SEQ ID NO:22) and 1MEL (nucleotide sequence SEQ ID NO:17, protein sequence SEQ ID NO:18).

[0025] b) Amino acid sequence (SEQ ID NO:24) and corresponding DNA sequence (SEQ ID NO:23) of a primary scaffold that is used to obtain optimal (secondary) scaffolds. Amino acid sequences were obtained using modeling software (Modeller 6.0 and Insight II) in combination with DNA and protein analysis software (Vector NTI suite7.0, InforMax). Light and heavy chain variable regions were used as a template to design this primary scaffold. The

numbers (1-9) indicate beta-elements and L1-L8 indicate loops similarly as described in earlier figures. Underlined regions (L2, L4 and L8) indicate affinity regions located at one site of the proteins. L2, L4 and L8 correspond to the location of CDR regions.

[0075] A primary scaffold is designed as a template to generate secondary scaffolds with improved properties or features by mutational strategies. A primary scaffold design has to meet some desired criteria. First, it should be highly soluble, very stable and easy to fold. The solubility of proteins can be changed, for example, by changing amino acids that have side chains that are exposed to the environment. Hydrophilic amino acids like lysine, aspartic acid, arginine, histidine, serine, threonine or glutamic acid can improve the solubility. The stability and folding properties can be improved by the addition of intra-molecular bonds. The most powerful bond is found in sulphide-bridges between two opposed cysteine residues. However, bacterial cells, in contrast to prokaryotic cells like yeast cells, can have problems with cysteine bridge formations and, therefore, only a limited number of cysteine bridges should be incorporated in proteins that have to be expressed in bacterial cells. Because the used scaffold is structurally based upon immunoglobular domains, and because at least some bacterial proteins contain such structures or deviations of these, one can incorporate the structural and amino acid sequence knowledge of these proteins into new scaffolds that are based upon these structures. Using Modeller 6.0, Insight II individual or stretches of amino acid changes were checked for stability prognoses in comparative modeling situations. Especially the outer amino acids of the structures were initially modeled. Vector NTI suite 7.0 software was used to design the corresponding DNA sequences including codon optimizations for *E.coli* cells and the incorporation of desired restriction sites. Variable regions from human, mouse and lama (especially VHH) have been used as a template to design a new scaffold. This scaffold therefore contains 9 beta-elements (see, figures 1 and 2) arranged in 2 beta-sheets. The four even-numbered loops can form a binding region on one side of the scaffold (for example, as found in antibodies) while the odd-numbered loops on the other side of the scaffold can be used for a binding region for another or even the same molecule. From all models that were created using comparative modeling software, the following primary scaffold with 9 beta-elements was chosen

based on energy minimization, potential folding properties, potential solubility and potential stability (nucleotide sequences of SEQ ID NO:1 and SEQ ID NO:23 connected by (NNN)X (NNN)_x; amino acid sequences of SEQ ID NO:2 and SEQ ID NO:4 connected by (Xaa)_x):

[0081] The MCS of vector pBluescript KS+ (Stratagene) is amplified with the T7 (5' AATACGACTCACTATAG 3' (SEQ ID NO:65)) and the T3 (5' ATTAACCCTCACTAAAG 3' (SEQ ID NO:76)) primer(Isogen) in the following polymerase chain reaction (PCR): 50 ng of vector DNA is amplified by Taq DNA polymerase (0.3 units per reaction, Promega) in a 50 ml reaction containing 50 ng of T3 and T7 primer, 2 mM MgCl₂, 10mM Tris-HCl (pH 8.3), 50 mM KCl, 200 μM dNTPs (Roche Diagnostics) and with or without 200 μM dITP (Boehringer). After 30 cycles (45 sec 94 °C, 45 sec 44 °C and 45 sec 72 °C) (Primus 96^{plus}, MWG-Biotech), a part of the PCR products is digested with the endonucleases XbaI, PstI, EcoRI, Xho I (Gibco BRL; 20 units per reaction) and run on a 2.5 % agarose gel, while the rest of the PCR products are cloned into the vector pGEMTeasy (Promega) and sequenced. The PCR products produced in the absence of dITP are digested completely, while the PCR products produced in the presence of dITP are only partly digested, showing mutagenization of the recognition site of the endonuclease.